

LA-UR-21-24987

Approved for public release; distribution is unlimited.

Title: (U) Archiving Physics Data for Today and Tomorrow

Author(s): Ham, Michael Isaiah

Intended for: Physics Cafe Presentation

Issued: 2021-05-24

Disclaimer:

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Triad National Security, LLC for the National Nuclear Security Administration of U.S. Department of Energy under contract 89233218CNA000001. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

Welcome to Physics Café



Mike Ham, P-2:

(U) Archiving Physics Data for Today and Tomorrow

April 15, 2021

My Data Background

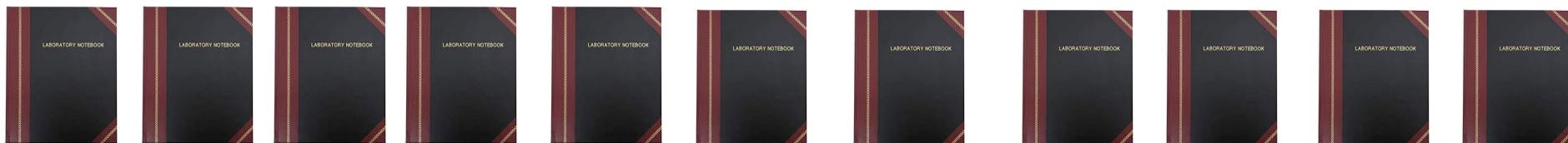


Project	Types of Data Collected / Accessible to community	Status of Data
Legos	Mental / No	Still in use
UNT Liquid Nitrogen Car	Notebooks / No	Discarded during a move
UNT Accelerator Laboratory	Notebooks / No	Publications
UNT Neurophysiology Laboratory / LANL T Division	<ol style="list-style-type: none"> 1. Notebooks / No 2. Electronic Database / No 3. Electronic data files / No 4. Electronic Documents / Only Published 	<ol style="list-style-type: none"> 1. Publications 2. Removable hard drives that maybe work (13 years)
LANL Computer Vision	<ol style="list-style-type: none"> 1. Notebooks / Yes 2. Myriad of computer programs and data files / No 3. Electronic Documents / Only Published 	<ol style="list-style-type: none"> 1. Notebooks in my office 2. Computer programs may exist on a thumb drive 3. Publications 4. Data likely inaccessible (computer lockout)
LANL Reaction History	<ol style="list-style-type: none"> 1. Notebooks / Yes 2. Electronic Documents / Yes 3. Scanned Repository / Yes 	<ol style="list-style-type: none"> 1. X Online Vault 2. Red Repository 3. SSDB 4. Common Modeling Framework



LANL did/does some things well with data

- **Most high level paper documents from historic shots were preserved.**
 - Most results data were preserved (films)
 - We have an active archival effort (WRS)
 - Culture of recording data
- **Culture has reverence for data when answering important problems**
- **When I got to LANL we were told to preserve notebooks**
 - Notebooks are provided! (your own personal repository)
 - *Have saved multiple experiments, use a notebook*
 - Probably mostly useless for others to read, but that's okay



LANL does some things badly with data



- Some historic data was deemed unimportant and discarded after testing
 - Cables
- Archival efforts tend to be extremely ad-hoc
- Institutional data repositories are not well known and not user friendly
 - No institutionally accepted way to archive your digital data for use by the laboratory, it's up to you and your team/group/division.
 - That has some advantages (flexibility)
 - But also disadvantages in that the data are everywhere and nowhere

LANL Inside

Computing Employees Environment Finance News Safety Science Security Services

Data & Servers » Central Storage » Project Storage

Project Storage

Description of service

Shared disk space where a group of people working on a project can store project data. Only individuals involved in the project have access to the data.

▼ Standard service features

- Data is backed up and recoverable
- Approved project storage spaces are institutionally funded.
- Projects spaces >10TB require validation prior to provisioning.
- Past versions of data available as snapshots at defined intervals - allows multiple reversions up to 6 months
- Remote set-up available
- Support available Monday through Friday, 8 a.m. - 5 p.m.

▼ User requirements

- DCStorage customer (self-supported organizations can arrange for this service)
- Project's funding official must fill out [DCStorage Shared Space Contract \(pdf\)](#)
 - Please download the contract to your machine, fill it out, and attach it to your AskIT (Remedy) ticket
 - Must be reviewed by project's DCStorage technician prior to submission

Service self-help

► FAQs

► Additional resources

askIT? call 5-4444 submit chat request

IT Internal Requests

FAQs

Additional Resources

- Data access
- Data stability and recovery
- Data storage

<https://int.lanl.gov/computing/data-servers/central-storage/storage-types.shtml>

Making data available to the future



Home Getting Started FAQs Create New Document SS-Production What's New Investment to generate HE data: \$12.8M Welcome Michael I. Ham My Preferences Logout

Small Scale Experiment Database

Identifier (Shot#) Full-text

Database Views

- Browse All Documents
- Advanced Search
- By Author
- By HE Material Type
- By Identifier
- By Tags

Welcome to the Small Scale Experiment Database. Today is Wed Apr 14, 2021 Refresh Page

My Drafts (1)

View all documents created by me Create a new experiment document

Exp num	Title	Last revised	Category	Status
BLUEROOM-1	Blue Room Experiment	3/11/21	General Experiment	In Process -- Michael I. Ham -- 9/28/20 10:03 AM

What's New (0)

ATTACHMENTS

Description of Attachments: gamma_PSR_BlueRoom_2020.tar.gz are the files collected by the four pack gamma detector and the medusa detector

Ch1: trigger
Ch2: Mach Zehnder # 2 connected to PMT A
Ch3: Mach Zehnder # 1 connected to PMT B
Ch4: Cable connected to PMT A
Ch5: Cable connected to PMT B
Ch6: Cable connected to PMT C
Ch7: Cable connected to PMT D
Ch8: Medusa PMT

All Attachment(s):

Download All Attachments (approximate size: 227 MB; estimated zip time: 96 seconds)

File Type	File Name	Size (KB)	Last Modified
Other	MSO_0_PSRData_BlueRoom_2020.tar.gz	64928	12/11/20 12:10 PM
Other	MSO_1_PSRData_BlueRoom_2020.tar.gz	110171	12/11/20 12:10 PM
Other	gamma_PSR_BlueRoom_2020.tar.gz	57084	12/1/20 2:34 PM

<http://smallscale-y.lanl.gov>

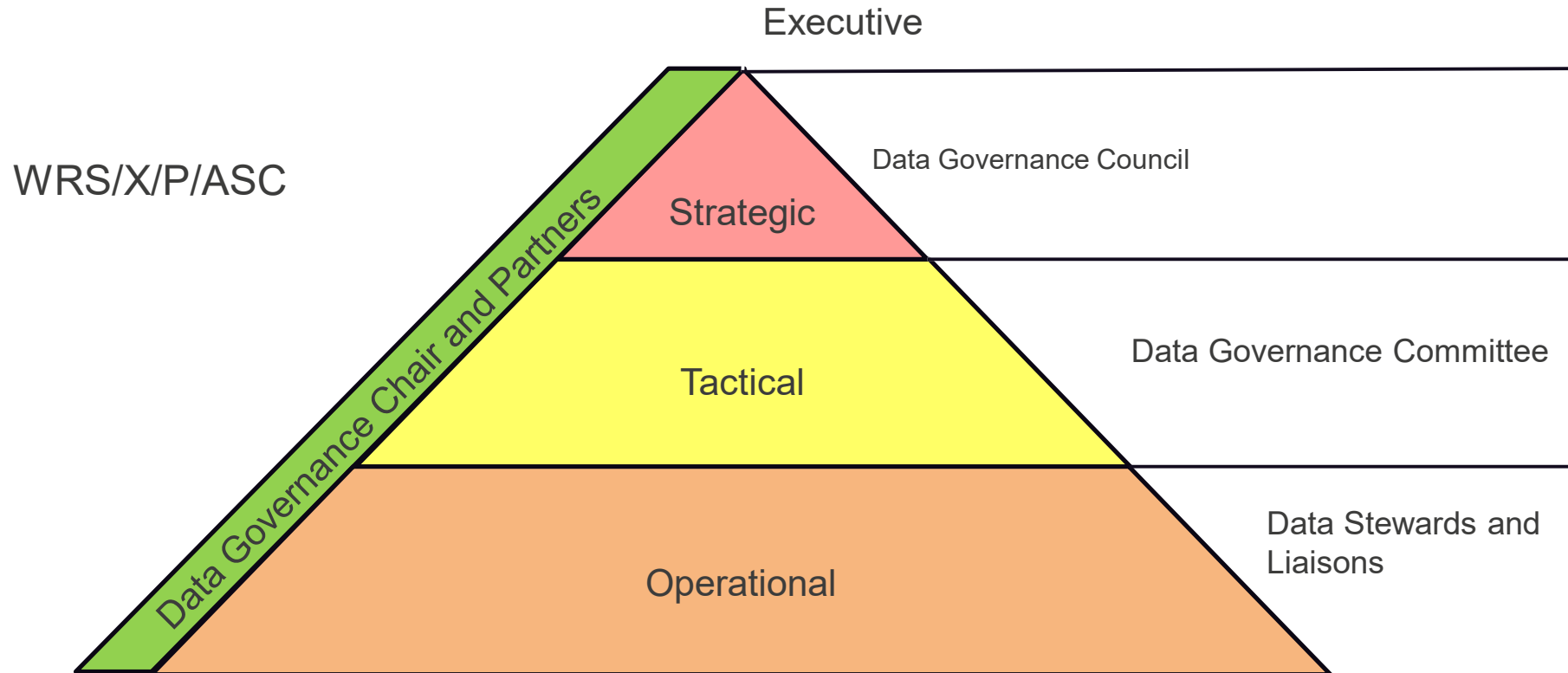
What am I Currently doing?



- **Assisting in updating internal digital archives**
- **Storing Data in the Small Scale Data Base (SSDB)**
 - Blue Room data (yellow)
 - MIE scattering data (yellow)
 - UGT data (red)
 - Surface Imaging Data?
- **Working directly with X-division to put data with pedigree into the Common Modelling Framework (CMF)**
 - Not complete agreement that this is where it should go, but it does work and is simple to implement.
- **Assisting in Establishing Data Governance Body with WRS**

Data Governance Body

Roles and Responsibilities



Roles and Responsibilities

